AD_____

Award Number: W81XWH-06-1-0100

TITLE: Affinity-Based Serum Proteomics for Ovarian Cancer Early Diagnosis

PRINCIPAL INVESTIGATOR: Martin W. McIntosh, Ph.D.

CONTRACTING ORGANIZATION: Fred Hutchinson Cancer Research Center
                          Seattle, WA 98109-1024

REPORT DATE: December 2007

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
        Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
                        Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
01-12-2007	Annual	15 NOV 2006 - 14 NOV 2007

4. TITLE AND SUBTITLE

Affinity-Based Serum Proteomics for Ovarian Cancer Early Diagnosis

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-06-1-0100

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Martin W. McIntosh, Ph.D.

E-Mail: mmcintos@fhcrc.org

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Fred Hutchinson Cancer Research Center
Seattle, WA 98109-1024

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Our research project is intended to exploit unique characteristics of phage and yeast recombinant antibodies as the basis for a serum biomarker discovery platform for ovarian cancer. In brief, we select from large recombinant libraries those binding sequences which bind to cancer related material but not to control serum, then we evaluate these sub libraries in high throughput using novel recombinant antibody arrays probed with serum from our serum repository. At present, we are on track based on our initial proposal. We have (1) selected a well-balanced group of cases (serum and proximal fluid) and controls for our initial discovery, (2) identified thousands of unique binding sequences that bind to the cases and not controls, (3) printed over 1,700 recombinant antibodies on high density arrays and (4) probed those arrays with individual sera from 50 cases (including early and late stage, and high and average risk women) and 50 asymptomatic controls. In addition to these tasks, we have also undertaken several research tasks to further optimize our experimental protocols. These include a series of shotgun proteomics experiments used to characterize the protein constituents of the clinical materials used in our selection, an evaluation of multiple array normalization and processing protocols to tailor data analysis to our array platform, and improved methods for high throughput shuffling (yeast library only) and purification of antibodies. At present, materials from our project include libraries of binding agents and data, including microarrays profiling dozens of specimens and mass spectrometry data characterizing the constituents of ovary tumor proximal fluid. To date, the major findings of our proposal include the proof of principle that (based on our data analysis) the panning and array procedures are capable of evaluating thousands of unique antibodies, and that (based on the proteomics measurements) the selection material is rich in putative biomarkers.

15. SUBJECT TERMS
No subject terms provided.

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON USAMRMC
a. REPORT	b. ABSTRACT	c. THIS PAGE	UU	10	19b. TELEPHONE NUMBER (include area code)
U	U	U			

Table of Contents

Affinity-Based Serum Proteomics for Ovarian Cancer Early Detection
Martin McIntosh, PhD, Principal Investigator

INTRODUCTION:

Our research project is intended to exploit unique characteristics of phage and yeast recombinant antibodies as the basis for a serum biomarker discovery platform for ovarian cancer. Our initial plan intended to select from millions of recombinant antibodies those thousands of binding sequences which bind to cancer serum but not to control serum, then evaluate these sub libraries solely using high throughput recombinant antibody arrays, probed with serum from our serum repository.

At the start of last year we had (1) selected a well-balanced group of cases (serum and proximal fluid) and controls for our initial discovery, (2) identified thousands of unique binding sequences that bind to the case pools and not controls, (3) have printed over 1,700 recombinant antibodies on high density arrays and (4) probed those arrays with individual sera from 50 cases (including early stage and late stage, and high and average risk women) and 50 asymptomatic controls.

Since last year we have also:

1) Found that during our selection process, selecting scFv which bind to cancer ascites (fluid proximal to the tumor) rather than first selecting those which bind to serum, yields higher quality classification biomarkers than selecting based solely on plasma alone. Thus, over much of the last year we have made exclusive use of these proximal fluid-based scFv in our profiling.

2) Developed a pipeline for evaluating the library using quantitative mass spectrometry methodologies and have these methods in place for evaluating our novel scFv libraries. See results in Aim 3 below, and which is also summarized in a manuscript which is currently under revision (Scholler, et al) .

3) Added to the array a total of 1080 full-length antibodies to the array, many from putative ovarian cancer markers selected from our proximal fluid selection, as described above.

4) Probed and analyzed data from an additional 115 and 96 ovarian cancer case arrays and control arrays in two batches, exceeding the throughput initially planned.

5) Published a validation report of the array format (Loch, et al 2007), and have established that markers contained in the hypoxic response and Map kinase pathways may be potential sources of biomarkers.

6) Identified a potentially novel protein biomarker, PEBP1, using both arrays and mass spectrometry-based approaches.

7) Begun developing a scFv-based sandwich ELISA assay for PEBP1 to support further high throughput evaluation.

With progress over the past year, we are on track based on our initial proposal.

BODY:

We first relate our progress directly to the stated aims of our proposal, as a large part of our progress has followed the specific tasks and timeline initially proposed. In addition, we have made advancements in our protocols and procedures to make our experimental platform more efficient. We describe some, but not all, of those below.

Our initial aims are in italics, and our response is in normal text. We also divide the response, for the reader's convenience, into progress made overall through one year ago, and progress made through the previous one year covered by this report.

Aim 1 Tasks: (Months 1 to 6): To identify thousands of candidate ovarian cancer biomarkers from phage and yeast recombinant libraries by selecting sub-libraries against reactivity to common abundant proteins and a heterogeneous pool of control sera and for reactivity to a biological material (sera and/or tumor constituent proteins) from a heterogeneous pool of cancer sera.

1. Select cases and controls to be pooled and used for biomarker discovery panning.
2. Perform panning to enrich for scFv that bind constituents in cancer material

Summary of progress through December 2006 (see previous report for details):

- Selected scFv using sera and proximal fluid resulting in nearly 5,000 scFv which appear to bind to proximal fluid but do not bind to normal control sera.

- Developed procedure for high throughput shuffling of the yeast scFv sub-libraries and purify selected phage scFv shuffled 1700 scFv used in our array fabrication in Aim 2.

- Purified large volumes of scFv sufficient for printing on arrays.

- Constructed a small pilot array using scFv and monoclonal antibodies and compared results to ELISA methods.

- Performed mass spectrometry interrogation of the proximal fluid library to characterize its contents. Protein candidate list reported in previous report.

Progress since December 2006:

- Aim was completed in previous year.

Aim 2 Tasks (months 6 to 18): To profile serum from a heterogeneous set of 75 ovarian cancer cases and 75 matched controls using antibody microarrays containing the recombinant antibody sub-libraries.

1. Select cases and controls for profiling as described in Methods section.
2. Perform case-control analysis of the 75 cases and controls including quality assurance tests (by profiling arrays using the serum from the pools used to produce the sub-libraries)

Summary of progress through December 2006.

- Pilot data characterizing the quality of the scFv sub-library using sequencing at DNA level to confirm the diversity of the library.

- Pilot LC-MS/MS experiments to characterize the quality of the library and protein identifications. We established that the scFv used to pull-down proteins from serum contains a larger number of proteins also found in the proximal fluid, establishing other evidence that our proximal fluid-based approach is promising.

- We established our methods and procedures to normalize and evaluate antibody micro-arrays. Details are provided in previous report, and some summaries are given in our recent manuscript (Loch, et al 2007). In brief, (1) Correct background using "normexp" method, (2) Detect and filter out poor quality antibodies on a slide using measurements from multiple spots (3) Perform slide-dependent non-linear normalization using only good quality antibodies.

Summary of progress since December 2006.

In our experimental plan we proposed to evaluate our arrays by probing 75 cases, 75 controls. We have met and exceeded this goal, evaluating a far larger number of arrays, including 115 cases and 96 control (not including QA and QC arrays). We probed plasma in two sets of experiments, including:

Array set 1: ScFv from initial rounds of selection.

Samples: 42 controls (22 healthy; 10 surgical normal, 10 benign), 42 cases (36 late stage; 6 early stage)

Antibodies: Phage ScFv: 4608 (2688 were selected against cystic fluid, 1248 selected against cancer serum, 672 against other cancer types to help measure specificity), Yeast ScFv: 672 (all were selected ovary proximal fluid), Full Length: 384

Array set 2: ScFv from array set 1 which exceeded background.

Samples: 73 controls (47 healthy; 13 benign; 13 surgical normal), 46 cancers (8 early stages, 38 late stages)

Antibodies: Phage ScFv: 1824 (381 ovary: 261 were selected from cystic fluid; 117 were selected from cancer serum; 1,100 from other cancer libraries to evaluate specificity), Yeast ScFv: 192 (selected from proximal fluid). Full Length: 1144

Results from detailed analysis of the first round of array printing: Using our first batch of arrays we have demonstrated the overall validity of this platform to profile the human serum proteome using the full length antibodies (Loch, et al 2007). In summary, the array contained 320 full-length antibodies (monoclonal or polyclonal), each printed in triplicate. Arrays were probed with serum which included 31 serous ovarian cancer cases (a subset of all cases probed) and 34 matched controls. The antibodies were pre-selected to represent three groups:

Group 1 contained 12 antibodies to three previously validated biomarkers including CA125 (n=8; Bast, et al. 1981) HE4 (n=2; also known as WFDC2; Hellstrom, et al. 2003), and mesothelin (n=2; also known as SMR; McIntosh, et al. 2004).

Group 2 contained a total of 38 candidate biomarkers in need of further validation that were identified in our previous discovery studies or in the literature (Biade, et al. 2006; Bratt 2000; Davidson, et al. 2006; Frank and Carter 2004; Lau and Chiu 2007; Lim, et al. 2007; Liu, et al. 2006; Moubayed, et al. 2007; Treiber, et al. 2006; Witton, et al. 2003).

Group 3 was a discovery set of 270 antibodies to cytokines, angiogenic factors, cancer antigens, differentiation markers, oncoproteins, and signaling molecules, none of which had *a priori* expectations of being ovarian cancer biomarkers.

A total of 90 antibodies from this third group were also pre-specified to be one of three subgroups of interest, including 19 regulated by hypoxia, 61 that are part of the mitogen-activated protein kinase (MAPK) pathway, and 10 related to the phosphatidyl inositol 3-kinase (PI3K) pathway.

As expected, Group 1 antibodies performed the best, followed by groups 2 and 3. In addition to validating individual antibodies, we were able to establish that the sub-group of hypoxic response and MAPK proteins might as a group (Subramanian, et al. 2005) contain a rich source of biomarkers. These findings demonstrate that the performance of previously validated antibodies is recapitulated on this platform.

We also used this platform to evaluate a number of other putative biomarker candidates. This second group of antibodies, to biomarker candidates in need of validation, was overall found to be ranked higher compared to the remainder of the array (p<0.001, Wilcoxon-rank sum test) with an average ranking of 86.1 out of 320. Thus suggests that this group of antibodies contains one or more targets with the ability to classify case or controls, many of which we may pursue in the final aims of our project.

Other activities in support of this aim and the project:

Because the yeast scFv have been modified to permit in-vivo biotinylation (which we term biobodies or Bbs) (Bergan, et al). We have been attempting to develop streptavidin-coated slides (Xenopore) as a platform for microarrays using *in vivo*-biotinylated yeast recombinant antibodies Bbs work well in ELISA on streptavidin-coated plates so we hoped that they would adapt well to the microarray format, and improve our ability to identify low abundant protein differences.

We have also recently completed ranking a large number of our full length and scFv identified from the two sets of array analyses described above, and several interesting results are evident. In terms of the recombinant scFv, many performed very well at distinguishing case from control in both arrays. Of the top 15 scFv, the selections we performed on the cystic fluid account for 12 while the other 3 are from selection with serum. We are currently working on identifying the protein biomarkers that bind to these top scFv. We are accomplishing this using HPLC separation of serum lacking the top 20 abundant proteins into approximately 96 fractions followed by dot blotting each fraction to test for reaction with the highly performing scFv. Identification of the specific fraction then allows us to spread the gradient in that area to accomplish better separation. The positive fractions are then subjected to LC-MS/MS (Aim 3). Possible identifications will then be tested using commercial antibodies if they are available. We have also confirmed that a number of full-length antibodies predicted the presence of ovarian cancer very well. As expected CA-125 and Mesothelin performed well but in addition, MEK kinase-3, HIF-3α and HIF-1α were essentially equally good.

Aim 3: Overview: Independent validation of tumor biomarker candidates. Because some of the scFv found from Aim 2 may be to the same protein, to more specifically target our development to maximize biomarker yield, we will first identify (sequence) the top 100 targets using immunoprecipitation and tandem mass spectrometry. It is our goal to begin with the top performing biomarkers and sequence until 25 unique biomarker candidates have been identified.

Summary of progress prior to December 2006:

- No progress in the previous year.

Progress since December 2006:

A core part of our scFv identification scheme includes the use of mass spectrometry as a tool for identifying the protein target of the scFv. We have been investigating the efficiency of performing this

identification individually or also using larger libraries of scFv. The latter approach has led us to begin using a novel discovery approach to evaluate the scFv based on serum immunoprecipitation.

Briefly, and also summarized in Scholler et al (under revison) using cancer-specific *in vivo* biotinylated recombinant derived from the differentially selected yeast-display scFv, and analysis of the eluted serum proteins by label-free quantitative mass spectrometry. We used immunoprecipitation with a library of *in vivo-* biotinylated yeast recombinant antibodies, biobodies or Bbs. We compared IPs from serum from cancer cases to control serum using a Bb library selected for binding to cancer serum and depleted for those that bind normal serum. The Bbs and bound proteins were collected on streptavidin-coupled magnetic beads.

Analysis using shotgun mass spectrometry and label-free quantitation found a larger number of proteins captured by the scFv, including 38 yeast proteins and 54 bovine proteins, each of which are present from as contaminants or the media. In addition, 278 sequences of human origin were identified. The three most differential proteins include catabolic fragments of complement factors, EMILIN2, Von Willebrand factor and phosphatidylethanolamine-binding protein 1 (PEBP1).

To support this workflow, we further attempted to establish that the proteins were indeed ovarian cancer associated by acquiring existing antibodies to PEBP1 by performing ELISA assays and antibody microarrays. Specifically, we added these antibodies to our array and also used the antibodies to probe 30 ascites samples. PEBP1 was detected in 29 out of 30 ascites samples investigated and could discriminate ovarian cancer sera from controls (p=0.02) on the array. Finally, based on other funding sources, we have interrogated plasma samples using mass spectrometry-based approaches developed in our group and were able to confirm the elevation of PEBP1 in plasma of ovarian cancer cases compared to controls.

We conclude that an extracytoplasmic form of PEBP1 is an ovarian carcinoma serum biomarker, and we are now evaluating several other markers we have identified, and generating larger numbers of mass spectrometry experiments to confirm the identify of the markers.

Aim 4: Overview: Repeat to target at risk groups and early detection biomarkers. It is our ultimate goal to identify markers that detect all cancer prior to their symptomatic development, and in this aim we intend to perform another round of biomarker discovery (panning) to maximize that possibility. Specifically, this time we will pan using individual sera from Aim 3 who were missed by the current panel from Aim 3, and also pan using the pre-clinical sera available to us; we will include in panning the latest collected pre-clinical sample that was not identified by the panel in Aim 3. If we also find that the marker panel is performing worse in some identifiable subgroups such as women with mutations in BRCA, we perform a round of selection using that misidentified subgroup. We then repeat aims 2 through 3 with these new sub-libraries.

Progress prior to December 2006:

- None, aim had not yet begun.

 Plans for this aim:

We have recently begun this aim but do not have any progress to report as this will be our emphasis (along with the identification aims of Aim 2) in the final year of our project.

KEY RESEARCH ACCOMPLISHMENTS:

We have completed a manuscript and systematic study of our platforms to establish their reproducibility and validity. We have also established PEBP1 as a putative marker for ovarian cancer in ascites and

plasma, and which results from our discovery platform. We have a larger list of markers which we are now pursuing.

REPORTABLE OUTCOMES:

Here we summarize reportable outcomes, summarized from the description above, in two areas, including Data and analysis, Affinity agents.

Data and analysis methods: The following data resources have been generated from the current project.
- Tandem MS data: We have generated several data sets which summarize the protein identifications from the scFv, and as a library, profiling the ovarian cancer proximal fluid. A total of 25 LC-MS/MS experiments were performed and the data are available for mining.

- Tandem mass spectrometry data summarizing the content of the precipitate of plasma.

- Array data: We have profiles of over 200 ovarian cancer cases and controls of various histologies and risk groups. These data are presently available for mining.

Affinity agents and arrays: The following affinity agents have been obtained and are stored in our freezers.
- We have selected over 5,000 scFv by panning what may be putative biomarkers for ovarian cancer.
- We have shuffled and purified over 1,000 of these scFv and these are available for high throughput evaluation.
- Putative binding sequences (biomarkers): We have identified a subset of 20 scFv which are putative binding sequences which may be useful biomarkers for cancer. We are in the process of confirming these markers and, once confirmed, their antigens (protein sequence) will be identified using LC-MS/MS.

PUBLICATIONS:

1. Use of high density antibody arrays to validate and discover cancer serum biomarkers. Molecular Oncology, Volume 1, Issue 3, December 2007, Pages 313-320. Christian M. Loch, Arturo B. Ramirez, Yan Liu, Cassandra L. Sather, Jeffrey J. Delrow, Nathalie Scholler, Barbara M. Garvik, Nicole D. Urban, Martin W. McIntosh and Paul D. Lampe.

Manuscripts under revision:

 2. Use of cancer-specific in vivo biotinylated recombinant antibodies for serum biomarker discovery: Identification of PEBP1 as a candidate ovary cancer biomarker, Nathalie Scholler, Jennifer A. Gross, Barbara Garvik, Lance Wells, Yan Liu, Christian M. Loch, Arturo B. Ramirez, Martin W. McIntosh, Paul D. Lampe, and Nicole Urban.

CONCLUSION:

We have established that our approach to using proximal fluid to target biomarkers is working highly efficiently, and will be our source of biomarkers moving forward. We have also established that we are able to identify biomarkers using both mass spectrometry and micro-arrays. We have identified several candidate markers which we will be evaluating over the next year.

REFERENCES:

Loch, C.M., Ramirez, A.B. , Liu, Y., Sather, C.L., Delrow, J.J., Scholler, N., Garvik, B.M., Urban, N.D., McIntosh, M.W., and Lampe, P.D. Use of High Density Antibody Arrays to Validate and Discover Cancer Serum Biomarkers. (2007) Molecular Oncology, 1, 313-320.

L. Bergan, J. Gross, B. Nevin, N. Urban, N. Scholler. Development and in vitro validation of anti-mesothelin biobodies that prevent CA125/Mesothelin-dependent cell attachment. Cancer Letters, Volume 255, Issue 2, Pages 263-274 Epub 2007 Jun 7.

APPENDICES:

No appendices are provided.

SUPPORTING DATA:

Raw and processed MS/MS data are available on request.